THE ASTROPHYSICAL JOURNAL, 503:923–937, 1998 August 20
© 1998. The American Astronomical Society. All rights reserved. Printed in U.S.A.

EVOLUTION OF THE SOLAR NEBULA. IV. GIANT GASEOUS PROTOPLANET FORMATION

ALAN P. BOSS
Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, NW, Washington, DC 20015-1305;
boss@dtm.ciw.edu

Received 1998 February 3; accepted 1998 March 27

ABSTRACT

The discovery of the first extrasolar planets, with masses in the range of $\sim 0.5\ M_{\rm Jup}$ ($M_{\rm Jup}$ = Jupiter mass) to $\sim 3\ M_{\rm Jup}$, demands a reevaluation of theoretical mechanisms for giant planet formation. Here we consider a long-discarded mechanism, forming giant planets through the gravitational instability of a protoplanetary disk. Radiative hydrodynamical calculations of the thermal structure of an axisymmetric protoplanetary disk with a mass of $\sim 0.13\ M_\odot$ (inside 10 AU), orbiting a solar-mass star, predict that the outer disk may be cool enough ($\sim 100 \pm 50$ K) to become gravitationally unstable. This possibility is investigated here with a fully three-dimensional hydrodynamics code. Growth of significant nonaxisymmetry occurs within a few rotation periods of the outer disk and can result in the formation of several discrete, multiple-$M_{\rm Jup}$ clumps in $< 10^3$ yr. These giant gaseous protoplanets (GGPPs) are gravitationally bound and tidally stable and so should eventually form giant planets. Modest-sized solid cores may form through dust grain growth and sedimentation prior to the centers of the GGPPs reaching planetary densities. The inner disk remains nearly axisymmetric throughout these phases, suggesting a scenario in which the formation of terrestrial planets occurs slowly through collisional accumulation in the hot inner nebula, while rapid formation of GGPPs occurs in the cooler regions of the nebula. Falling disk surface densities would restrict GGPP formation to an annulus, outside of which icy outer planets would have to form slowly through collisional accumulation. GGPP formation occurs for both locally isothermal and locally adiabatic disk thermodynamics, provided that the Toomre Q stability parameter indicates instability ($Q_{\rm min} \approx 1$). Low-order modes, especially $m = 1$ and 2, are dominant. Provided that a means can be found for inducing massive protoplanetary disks to undergo the GGPP instability (e.g., clumpy accretion of infalling gas onto a marginally stable disk), the GGPP mechanism appears to be a prompt alternative to the long-favored but protracted core accretion mechanism of giant planet formation. Observations hold the promise of deciding which of these two mechanisms is preferred by young stars.

Subject headings: accretion, accretion disks — hydrodynamics — solar system: formation — planetary systems

1. INTRODUCTION

The discovery of the first confirmed extrasolar planet (Mayor & Queloz 1995; Marcy et al. 1997) and the rash of subsequent detections (Butler & Marcy 1996; Gatewood 1996; Butler et al. 1997; Cochran et al. 1997; Noyes et al. 1997) have produced long-awaited information about the outcome of the planet formation process around other solar-type stars. Because these unseen companions have masses comparable to that of Jupiter ($\sim M_{\rm Jup}$) and nearly circular orbits (with one understandable exception—see Holman, Touma, & Tremaine 1997; Mazeh, Krymolowski, & Rosenfeld 1997), they are generally inferred to be gas giant planets similar to Jupiter and Saturn. Compared to the null results of just a few years ago (Walker et al. 1995), we now know that extrasolar giant planets do exist, at least around a fraction of nearby stars. A reexamination of the possible mechanisms for giant planet formation in the light of these and other recent advances thus seems timely.

The leading model of giant planet formation is core accretion (Mizuno 1980; Pollack 1984). Core accretion is predicated upon two distinct steps, (1) the formation of a roughly $\sim 10\ M_\oplus$ core and (2) the accretion of considerable nebular gas by the solid core. The $\sim 10\ M_\oplus$ core is assumed to form through the same process of collisional accumulation of planetesimals and planetary embryos that led to the formation of the terrestrial planets (see, e.g., Wetherill 1990). Several problems have become apparent with the core accretion model, with regard to timescales and core masses.

The time required to form a $\sim 10\ M_\oplus$ core is estimated to be on the order of 10^6 yr (Lissauer 1987), even in an optimistic scenario in which the disk surface density is increased between 5 AU and 10 AU, the gravitational enhancement of the collisional cross section for accumulation is maximized, and an infinite reservoir of feedstock is assumed to be available (Pollack et al. 1996). Recent models suggest that up to an additional $\sim 10^7$ yr is required for a $\sim 10\ M_\oplus$ core to accrete $\sim 300\ M_\oplus$ of nebular gas (Pollack et al. 1996). The total time thus falls at the upper end of the estimated lifetimes of circumstellar disks, $\sim 10^5$–10^7 yr (Strom, Edwards, & Skrutskie 1993). If the disk gas disappears prior to the formation of 10 $M_\oplus$ cores, the cores will be unable to accrete the H and He gas needed to form the envelopes of the gas giant planets like Jupiter and Saturn—instead, Uranuses and Neptunes would result.

Another problem for the core accretion model has come from a new generation of models of the interiors of Jupiter and Saturn (Chabrier et al. 1992; Guillot et al. 1994; Guillot, Gautier, & Hubbard 1997). Using improved equations of state for hydrogen and other refinements, it has been found that the observational constraints could be matched with considerably smaller core masses than had previously been thought to be necessary—between 0 and 12

$M_\oplus$ for Jupiter and between 1 and 13 $M_\oplus$ for Saturn. Previous estimates (see, e.g., Stevenson 1982) placed the core masses at between 10 and 30 $M_\oplus$ for Jupiter and between 15 and 25 $M_\oplus$ for Saturn. If the true core masses of Jupiter and Saturn lie at the lower end of these ranges, $\sim 1\ M_\oplus$, then the cores may not be massive enough to capture massive gaseous envelopes. Mizuno (1980) found the critical core mass to be $\sim 10\ M_\oplus$, independent of location in the disk, and this result was taken at the time as strong evidence for the core accretion model.

Additional difficulties may be raised as well. Weidenschilling (1997) found that even with optimistic assumptions for the core accretion model (Pollack et al. 1996), a detailed calculation of the first step (runaway accretion of planetesimals by a growing planetary embryo) showed that accretion took longer than expected and stalled with the formation of 2–3 $M_\oplus$ planetary cores, well below the desired mass of $\sim 10\ M_\oplus$.

Given these problems with the core accretion model, it seems prudent to consider other mechanisms for giant planet formation. One such model was proposed by Kuiper (1951) and advocated by Cameron (1978)—giant gaseous protoplanet (GGPP) formation. In the GGPP model, the solar nebula must be massive and cool enough to become gravitationally unstable. The gravitational instability is then envisioned to lead to the formation of self-gravitating clumps of gas and dust, called GGPPs. Because the instability occurs over a dynamic timescale (about an orbital period), GGPP formation can occur even during the abbreviated lifetime of the most ephemeral circumstellar disk.

The GGPP mechanism has been rejected historically because it seemed to be incapable of explaining the large, startlingly similar core masses of the giant and outer planets. The newly revised core masses for the giant planets have at least partially removed this objection, and the discovery of extrasolar planets with masses significantly greater than that of Jupiter suggests that a mechanism capable of forming massive giant planets is at work. However, a thorough investigation of GGPP formation has not yet been performed, so the mechanism has not been subjected to the same degree of scrutiny that has uncovered certain flaws in the core accretion model. The purpose of this paper is to initiate such an investigation, in order to discover the advantages and disadvantages of the GGPP mechanism.

Boss (1996) used a radiative hydrodynamics code to calculate a suite of axisymmetric (two-dimensional) solar nebula models with a common thermal characteristic: a relatively hot (1000 K) inner nebula, surrounded by a relatively cool (100 K) outer nebula. In the model with a disk mass of 0.13 $M_\odot$ (inside 10 AU), the surface density of the disk was close enough to the critical value for gravitational instability (Toomre 1964) that it appeared likely that such disks might be marginally gravitationally unstable in their outer regions. The $M_d = 0.13\ M_\odot$ model serves as the basis for the three-dimensional calculations of GGPP formation presented here.

2. NUMERICAL METHODS

The disk models were calculated with the three-dimensional hydrodynamics code described in detail by Boss & Myhill (1992). The code calculates finite-difference solutions of the equations of hydrodynamics and gravitation for a neutral, inviscid gas. The code was designed to be second-order accurate in both space and time through the addition of temporal corrections to the advection terms and a predictor-corrector treatment for body forces. The code was modified (as in Boss 1997a) to use the van Leer monotonic interpolation formula (van Leer 1977) to interpolate interface fluxes for both the density and the specific momenta or energy (the interface flux involves the product of these two quantities). In the original version of the code (Boss & Myhill 1992), only the interface density was obtained by van Leer interpolation, and the specific momenta and energy were obtained by averaging the adjacent cell center values. While the original code is formally more accurate, the modified code is considerably more stable numerically, permitting the use of significantly larger time steps (up to $\sim \frac{1}{2}$ of the Courant time step).

Spherical coordinates are used, with a radial grid ($N_r = 50$) that is uniformly spaced between 1 and 10 AU and a θ grid ($N_\theta = 23, \pi/2 \geq \theta \geq 0$) that is compressed into the midplane to provide adequate vertical resolution of the disk ($\Delta\theta = 0\overset{\circ}{.}3$ at the midplane). The ϕ grid is uniformly spaced with $N_\phi = 64$ for most models, leading to $\Delta\phi = 5\overset{\circ}{.}6$. In general, no symmetry is assumed in the ϕ variable, so that in a Fourier decomposition ($e^{im\phi}$) both even and odd m modes can grow. The Poisson equation for the gravitational potential is solved by a spherical harmonic (Y_{lm}) expansion including terms up to $l, m = 16$.

The inner edge of the disk occurs at 0.5 AU, the boundary between the disk and the central protostar. In order to prevent the growth of spurious noise in the inner disk without resorting to a punitively small time step, the inner disk was kept artificially axisymmetric by imposing $v_r = v_\theta = 0$ and $v_\phi \neq v_\phi(\phi)$. These "inner boundary conditions" permitted the use of time steps much larger than otherwise possible. In order to maintain numerical stability, it was also found to be necessary to damp $v_r > 0$ and prevent $v_\theta < 0$ (motions away from the protostar and the disk midplane, respectively) throughout the entire disk. The outer boundary conditions were defined on a constant volume sphere of radius 10 AU, where $v_r = 0$, and v_θ was forced to remain positive. The boundary density was also fixed at its initial value. This has the effect of adding a small amount of mass to the disk during the evolution (typically a few percent), because the mass that flows inward toward the protostar (which should lead to a lower boundary density) is then effectively replaced by new mass on the boundary. This was done in order to make the disk respond as if it were just the inner portion of a disk of considerably larger radial extent.

Considering the importance of the outer boundary conditions for the growth of nonaxisymmetry in disks (see, e.g., Adams, Ruden, & Shu 1989; Shu et al. 1990; Heemskerk, Papaloizou, & Savonije 1992; Ostriker, Shu, & Adams 1992), it is worthwhile to examine the adopted outer boundary conditions. Because $v_r = 0$ on the outer boundary, any radial momentum that is incident on the outer boundary from inside the disk is added to the boundary cell, where it immediately vanishes. Outward-moving radial momentum is thus absorbed on the outer boundary, and no momentum can enter the grid from outside the spherical volume. In order to have a reflecting outer boundary condition, outward-moving radial momentum in 1 time step would have to be reflected and converted to inward-moving radial momentum in the next time step, which is clearly not the case in the present models. When reflecting outer boundary conditions are used, as might be expected to be appropriate for a disk with a sharply defined outer edge, the $m = 1$ eccentric instability occurs even for disks in which the Toomre (1964) stability criterion is substantially exceeded, $Q > 1$ (Adams et al. 1989; Shu et al. 1990). When the outer boundary is assumed to be absorbing rather than reflecting, the $m = 1$ instability is suppressed (Adams et al. 1989; Shu et al. 1990). In contrast, the inner disk boundary conditions were found to have little effect on the instability. The use of absorbing outer boundary conditions in the present calculations is thus very conservative, in that any perturbation growth that occurs will not be dependent on the assumption of a sharp, reflecting disk edge.

The central protostar is represented as a gravitating point mass with $M_s = 1\ M_\odot$. For most models, the location of the protostar was fixed at the center of the spherical coordinate system. As a result of the growth of odd m modes, the location of the center of mass of the disk eventually departed from its initial position, though the departure is generally quite small compared to the radius of the inner edge of the disk (0.5 AU). In order to rigorously preserve the location of the center of mass of the combined protostar/disk system at the origin, some models allowed the protostar to wobble in response to the growth of odd m models in the disk (Tomley, Cassen, & Steiman-Cameron 1991). The wobbling protostar can then act back on the disk, serving as an additional source of nonaxisymmetry (Adams et al. 1989).

The two-dimensional disk model (Boss 1996) used to define the initial conditions for the present three-dimensional models required Eddington approximation radiative transfer to determine the disk temperature profile. However, the three-dimensional models cannot afford a similarly rigorous treatment of radiative transfer because of the need to evolve the models for many orbital periods. Hence while the three-dimensional models start with the disk temperature distribution determined by the two-dimensional radiative transfer code, the subsequent evolution of the disks is taken to be either locally isothermal or adiabatic. That is, the temperature at each numerical grid point is either held fixed (locally isothermal) or is updated after each time step according to the adiabatic prescription $T \propto \rho^{\gamma-1}$, where ρ is the local gas density and the gas pressure depends on the density as $p \propto \rho^{\gamma}$. The adiabatic exponent γ is 7/5 for molecular hydrogen.

The isothermal assumption ($\gamma = 1$) would be valid if any compressional energy produced by the growth of disk nonaxisymmetry is quickly radiated away. The opposite assumption is the adiabatic assumption, where the disk retains all of its compressional energy and is unable to cool by radiation. Realistic disk behaviors presumably lie somewhere in between the two extremes of $\gamma = 1$ and $\gamma = 7/5$.

Each three-dimensional model was evolved for $\sim 10^6$ time steps. Given the spatial resolution of the models (70, 451 active grid points), a typical model required several months to run on a dedicated DEC Alpha workstation.

3. INITIAL CONDITIONS

All of the three-dimensional models presented here were based on the same initially axisymmetric disk (Boss 1996) with a mass of 0.13 $M_\odot$ between 0.5 and 10 AU. The suspected protoplanetary disks inferred to be orbiting around T Tauri stars typically have total masses of ~ 0.02 $M_\odot$, while they range in mass from ~ 0.001 to ~ 1 $M_\odot$ (Beckwith et al. 1990). Most of this mass is inferred to lie at large radii, though because the innermost disks are optically thick at millimeter wavelengths the amount of mass in the inner regions may well be underestimated. As a consequence, it is unclear if any of the T Tauri disks contain as much as ~ 0.13 $M_\odot$ within ~ 10 AU, making the present models perhaps unrealistically massive for T Tauri stars. However, the young stellar objects that predate the T Tauri star phase presumably have much more massive disks—indeed, it is hard to avoid the formation of a massive disk from the collapse of a rotating molecular cloud core (see, e.g., Yorke, Bodenheimer, & Laughlin 1993). The present models might then apply better to earlier phases of protostellar evolution.

It is also worthwhile to compare the assumed disk mass with that obtained by restoring the solar system's planets to solar composition. Weidenschilling (1977) found the resulting "minimum mass" for the solar nebula to lie in the range 0.01–0.07 $M_\odot$, within a radius of about 36 AU. The amount of mass inside 10 AU ranged from ~ 0.005 to ~ 0.05 $M_\odot$. While these estimates are only lower bounds on the nebular mass, they are significantly lower than the 0.13 $M_\odot$ assumed in the present models. The applicability of the present models to the formation of our solar system then requires that the solar nebula experienced a phase with an inner disk mass at least 2–3 times that of the minimum mass nebula.

The initial radial density profile in the disk midplane is $\rho_m \propto r^{-3/2}$, chosen in part to minimize the contributions of gas pressure to non-Keplerian rotation (Boss 1993). If the disk had a vertical height that increased linearly with orbital radius, then the surface density profile would be $\sigma \propto r^{-1/2}$. While this is approximately true, the detailed vertical hydrostatic balance leads to a surface density that is not a simple power law, but that varies approximately as $\sigma \propto r^{-0.6}$ inside 5 AU and as $\sigma \propto r^{-1.3}$ outside 5 AU. Both of these are somewhat flatter than the $\sigma \propto r^{-3/2}$ profile derived by Weidenschilling (1977) and used by Beckwith et al. (1990). However, Cameron (1995) noted that $\sigma \propto r^{-1}$ characterizes steady viscous accretion disks with a constant α parameter. Lissauer (1987) suggested that the solar nebula may have had $\sigma \propto r^{-1/2}$ out to Jupiter's orbit but then dropped off faster than $\sigma \propto r^{-1}$ outside Saturn's orbit, a situation very similar to that of the present models. Unfortunately, interferometric observations of suspected protoplanetary disks are unable to constrain surface density profiles inside radii of ~ 50 AU, and even at larger radii, simple power-law disk models do not seem to fit the observations (Lay, Carlstrom, & Hills 1997). Theoretical calculations of presolar nebula collapse are not yet able to predict surface densities inside ~ 5 AU (Yorke et al. 1993). Given these uncertainties, the assumed $\sigma(r)$ of the present models seems quite acceptable.

The vertical density profile of the disk was calculated at each radius assuming hydrostatic equilibrium of an adiabatic gas. Outside the disk, the initial density falls off with radius as $\rho \propto r^{-3/2}$, and the residual cloud envelope is assumed to be infalling toward the disk midplane, as specified in Boss (1993, 1996). Mass was assumed to be accreting onto the disk from the residual presolar cloud at a rate of $\sim 10^{-6}$ $M_\odot$ yr^{-1}, as is indicated by observations of young stellar objects (Ohashi et al. 1996).

The three-dimensional models generally are started with a nonaxisymmetric perturbation to the two-dimensional

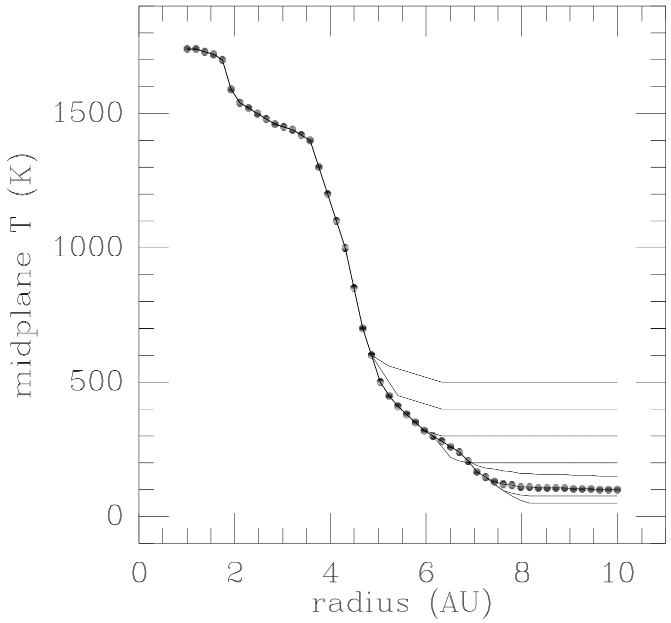

FIG. 1.—Midplane temperatures assumed as initial conditions for the three-dimensional models. Solid dots are the calculated T_m profile for the 0.13 $M_\odot$ disk from Boss (1996), leading to a temperature at 10 AU of $T_{10} = 100$ K. Solid lines show variations in the outer disk temperature profile employed in different three-dimensional models, with $T_{10} = 500$, 400, 300, 200, 150, 100, 75, and 50 K.

density distribution $\rho_{2D}(r, \theta)$ of the form

$$\rho_i(r, \theta, \phi) = \rho_{2D}\{1 + a_m[(\text{ran} + \cos(m\phi)]\} ,$$

where a_m is the amplitude of a perturbation of mode m (either $m = 1$ or 2, in general) and ran is a random number in the range [0, 1]. The amplitudes a_m were set equal to 0.01, whereas the amplitude of the noise resulting from these initial conditions was no more than ~ 0.001 for $m \leq 16$.

TABLE 1

INITIAL CONDITIONS AND RESULTS FOR $m = 2$ PERTURBATION MODELS

Model	γ	T_{10} (K)	$Q_{\min}$	Result	Number of GGPPs	t_f (yr)
m	1.0	50.	0.66	Unstable	3	497.
n	1.0	100.	0.94	Unstable	2	412.
o	1.0	150.	1.1	Marginal	2	656.
t	1.0	200.	1.3	Marginal	...	432.
p	1.0	300.	1.6	Stable	0	363.
r	1.0	400.	1.9	Stable	0	39.3
q	1.0	500.	2.1	Stable	0	27.3
x	1.1	50.	0.69	Unstable	...	37.6
v	1.1	100.	0.98	Unstable	2	518.
w	1.2	50.	0.72	Unstable	4	316.
u	1.2	100.	1.0	Unstable	2	464.
c	1.2	150.	1.3	Marginal	2	690.
d	1.2	200.	1.5	Stable	0	752.
y	1.4	50.	0.78	Unstable	2	520.
z	1.4	75.	0.96	Unstable	2	702.
s	1.4	100.	1.1	Marginal	...	566.

The two-dimensional models sought to determine the equilibrium structure of a disk that is being heated by infalling gas and self-contraction and cooled by radiative losses to interstellar space. The resulting midplane temperature profiles are depicted in Figure 1. The temperature plateaus in the inner disk around 1750 K and 1450 K are the result of thermal buffering caused by the loss of opacity due to water vapor and iron grains, respectively. A variety of astronomical, cosmochemical, and theoretical evidence supports the general trend exhibited by Figure 1 (Boss 1998a). However, because of the slow convergence toward an equilibrium solution in the outer disk, the temperature at 10 AU (T_{10}) is not well determined and is roughly $T_{10} = 100 \pm 50$ K on the basis of the two-dimensional models. Accordingly, T_{10} has been treated as a free parameter, in order to investigate the behavior of disks with varied thermal profiles. As shown in Figure 1, T_{10} has been varied in the range 50–500 K, with the higher values being intended primarily to demonstrate stability of the outer disk at sufficiently high temperatures. The initial temperature is 50 K in the infalling envelope for all the models.

Because the mass of the disk (0.13 $M_\odot$) is not negligible compared to that of the central protostar (1 $M_\odot$ for all models), the initial angular velocity had to be adjusted in order to start with the disks in approximate centrifugal balance. This was done by beginning with the disk having the angular velocity appropriate for Keplerian rotation around a protostar with a mass of 1.13 $M_\odot$. In addition, $v_r = 0$ and $v_\theta = 0$ inside the disk initially. This initial equilibrium was then tested by evolving the disk in two dimensions for many rotation periods, to ensure that the disk was indeed stable in the radial direction, and the initial angular velocity profile was then varied until a satisfactory result was obtained. The vertical equilibrium of the outer disk was less well preserved, however, but because of the decision to treat T_{10} a free parameter, such behavior was to be expected: e.g., the lower values of T_{10} lead to noticeable vertical contraction of the outer disk.

Varying T_{10} has the effect of varying the minimum value of the Toomre (1964) Q stability parameter in each disk. For an axisymmetric, thin, collisionless disk, Toomre (1964) showed that a nonaxisymmetric perturbation will be unstable if $Q < 1$. For a disk in Keplerian rotation, Q is defined as

$$Q = \frac{0.936 c_s \Omega}{\pi G \sigma} ,$$

TABLE 2

INITIAL CONDITIONS AND RESULTS FOR VARIED m PERTURBATION MODELS

Model	γ	T_{10} (K)	Protostar	Perturbation	Result	Number of GGPPs	t_f (yr)
o	1.0	150.	No wobble	$m = 2$	Marginal	...	656.
a	1.0	150.	No wobble	$m = 1$	Unstable	1	650.
aw	1.0	150.	Wobble	$m = 1$	Unstable	2	754.
ah	1.0	150.	Wobble	$m = 16$	Stable	0	544.
az	1.0	150.	Wobble	Axisymmetric	Stable	0	311.
aa	1.0	150.	Wobble	$m = 1, 2, 3, 4$	Unstable	2	675.
z	1.4	75.	No wobble	$m = 2$	Unstable	2	702.
b	1.4	75.	No wobble	$m = 1$	Unstable	1	596.

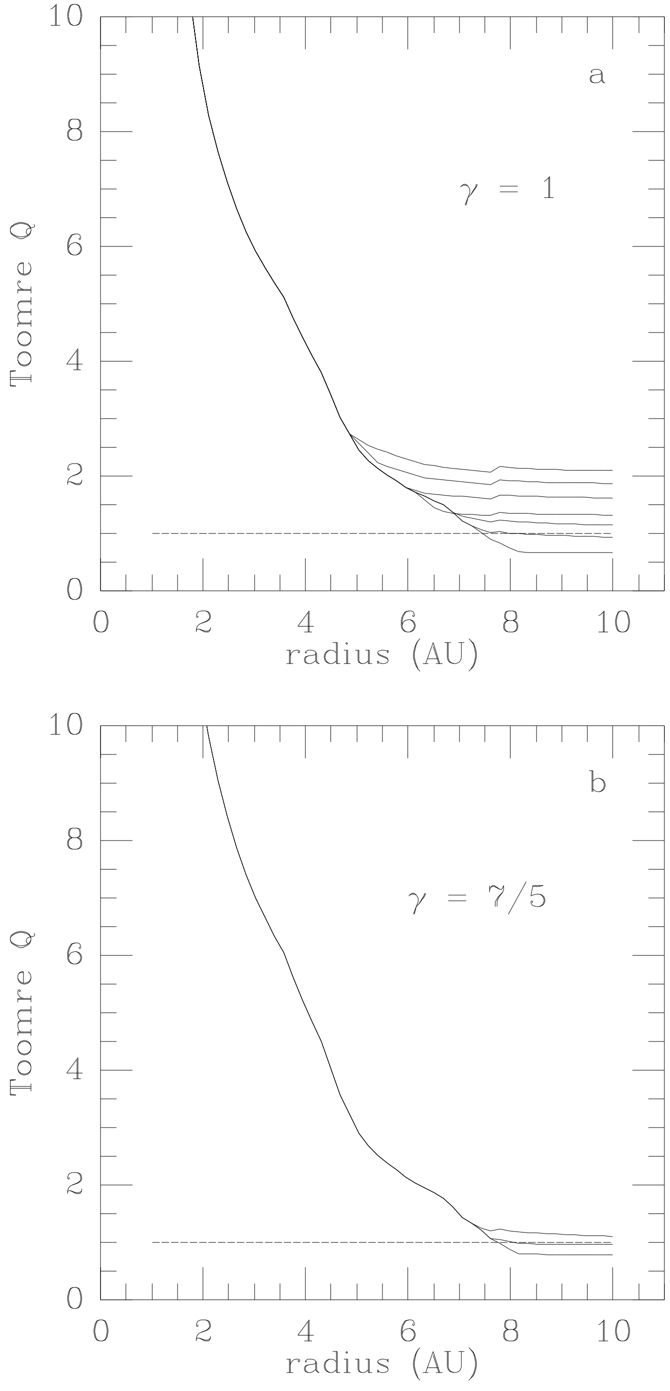

FIG. 2.—Toomre (1964) Q stability parameter as a function of radius for initial models with (*a*) $\gamma = 1$ ($T_{10} = 500$, 400, 300, 200, 150, 100, 50 K) and with (*b*) $\gamma = 7/5$ ($T_{10} = 100$, 75, and 50 K). Horizontal dotted line is the nominal threshold for instability ($Q < 1$).

where c_s is the sound speed, Ω is the angular velocity, and G is the gravitational constant. For locally adiabatic variations in the disk temperature, the sound speed is defined as

$$c_s = \left(\frac{\gamma p}{\rho}\right)^{1/2} .$$

Figure 2 depicts the radial dependence of Q for the three-dimensional models with varied T_{10} and varied γ—evidently the disks become more prone to instability in their cool outer regions.

Table 1 lists the initial conditions for the three-dimensional models that were given initial $m = 2$ density perturbations. Variations in both γ and T_{10} were explored systematically, in order to determine the critical values of Q for instability. The minimum initial value of $Q = Q_{\rm min}$ occurs at 10 AU in each model. Table 2 lists the initial conditions for three-dimensional models with variations in the form of the initial density perturbation and in the wobble of the central protostar. In addition to these models, several models in which the numerical resolution (N_ϕ) was varied in order to learn the effect on disk stability will be discussed.

4. RESULTS

4.1. *Classification of Outcomes*

Tables 1 and 2 list not only the initial conditions but also the basic outcome of each three-dimensional model at the final time (t_f). The outcomes are classified as one of three possibilities: stable, marginal, or unstable. Disks that were evolved far enough in time to form GGPPs are also classified by the number of distinct GGPPs evident at the final time.

The classification of outcomes is based upon the discussion of stability for nondissipative systems by Chandrasekhar (1961): when perturbed, stable systems undergo undamped oscillations with a characteristic frequency, while unstable systems experience exponential growth (with time) of the perturbation. Marginal systems are then those that fall between these two well-defined outcomes. It is of course much easier to demonstrate instability than stability—an apparently stable disk may become unstable at very large times. Practically speaking then, we can classify the disks only on the basis of their behavior during a fixed period of time, taken here to be about 13 rotation periods at 10 AU, or about 360 yr.

4.2. *Varied Outer Disk Temperatures*

Models p, o, and n resulted in three different behaviors. These three models constitute a sequence with fixed $\gamma = 1$ and varied $T_{10} = 300$, 150, and 100 K, respectively.

Figure 3 shows the time evolution of the $m = 2$ mode for model p, with $Q_{\rm min} = 1.6$, which is classified as a *stable* model. The degree of nonaxisymmetry clearly undergoes oscillatory behavior in the outer disk, with a period on the order of 60 yr near 10 AU, about twice the orbital period at 10 AU, $P_{10} = 28.5$ yr (at 1 AU, $P_1 = 0.95$ yr). The period of the oscillation increases with increasing orbital radius. The lower amplitude structure in the inner disk has a sawtooth appearance and disappears as the size of the time step is decreased and hence appears to be largely numerical noise. The very innermost disk is artificially forced to remain axisymmetric in order to avoid the very small time step that would be needed to prevent the growth of similar numerical noise inside a few AU. Considering that the calculation extends to ~13 rotation periods at 10 AU and to ~400 rotation periods at 1 AU, the degree of numerical stability evident in Figure 3 is remarkable.

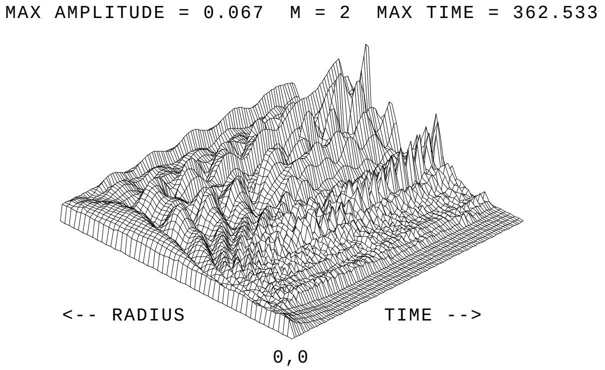

FIG. 3.—Time evolution of the amplitude of the $m = 2$ density mode as a function of disk radius in the stable disk model p ($Q_{\min} = 1.6$). Disk radius increases to left (spanning 1–10 AU) and time increases to right (0–363 yr). Starting from an initial amplitude of 0.01, the maximum amplitude rises to only 0.067 after many rotation periods, in an oscillatory manner, indicating the nonaxisymmetric stability of the disk. The sawtooth amplitude in the inner half of the disk at later times is numerical noise.

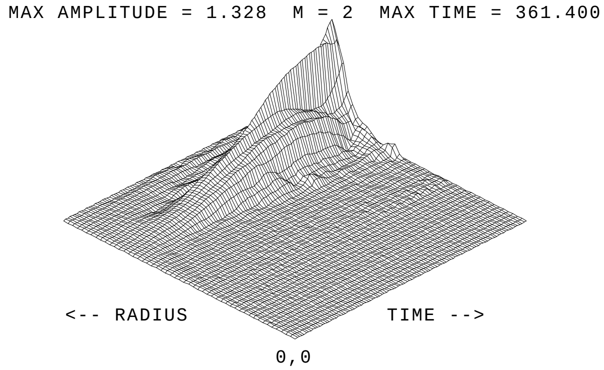

FIG. 5.—Time evolution of the amplitude of the $m = 2$ density mode as a function of disk radius in the unstable disk model n ($Q_{\min} = 0.94$), as in Fig. 3. The amplitude rises steadily around 8 AU and then saturates at a level greater than 1, signifying the formation of two GGPPs in the disk. The inner disk remains very nearly axisymmetric throughout, with negligible numerical noise.

Figure 4 shows the time evolution of the $m = 2$ mode for model o, with $Q_{\min} = 1.1$, which is classified as a *marginal* model. The $m = 2$ amplitude oscillates at first but then undergoes a period of exponential growth. Apparently marginally unstable disks will eventually fragment into GGPPs, though on longer timescales than unstable disks. By 656 yr, the maximum amplitude of the $m = 2$ model in model o had grown to 0.75, and the disk was well on its way to forming two GGPPs around 8 AU.

Finally, Figure 5 shows the time evolution of the $m = 2$ mode for model n, with $Q_{\min} = 0.94$, which is classified as an *unstable* model. The degree of nonaxisymmetry grows exponentially with time in the outer disk (Fig. 6), saturating at a level greater than 1. Even though GGPP formation has occurred in the outer disk, the inner disk remains very nearly axisymmetric.

This sequence of three models clearly shows the increasing tendency toward self-gravitational instability of the disk as its outer temperature is lowered from 300 K to 100 K (Fig. 1).

The stable disk models r and q (with $\gamma = 1$ and $T_{10} = 400$ and 500 K, respectively) were evolved over only a short period of time because of the absence of any significant growth during the first outer disk rotation period (maximum $m = 2$ amplitudes of 0.015 and 0.013, respectively). As could be expected, model m (with $T_{10} = 50$ K) was the most unstable of all and produced 4 GGPPs (two each at orbital radii of 7.5 and 9.5 AU) within 300 yr, when the $m = 2$ amplitude saturated at $a_2 \approx 2$. These four later merged into two, while a third GGPP formed around 5 AU because of the growth of the $m = 1$ mode.

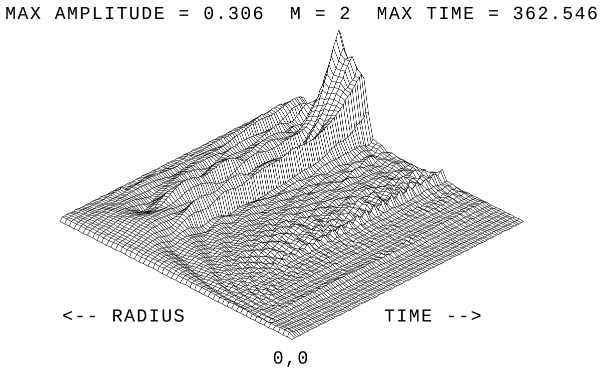

FIG. 4.—Time evolution of the amplitude of the $m = 2$ density mode as a function of disk radius in the marginal disk model o ($Q_{\min} = 1.1$), as in Fig. 3. The amplitude is initially oscillatory but then begins to grow rapidly. Some numerical noise is evident around 3 AU.

4.3. *An Unstable Disk*

Because the evolution of the gravitational instability is similar for all the unstable disks studied here, we present in some detail the results for a representative model.

The midplane evolution of model n is displayed in Figure 7. Starting from a nearly axisymmetric state ($a_2 = 0.01$, $a_{\pm 2} \approx 0.001$), the $m = 2$ mode begins to grow around a radius of 8 AU. Because of differential rotation, this non-

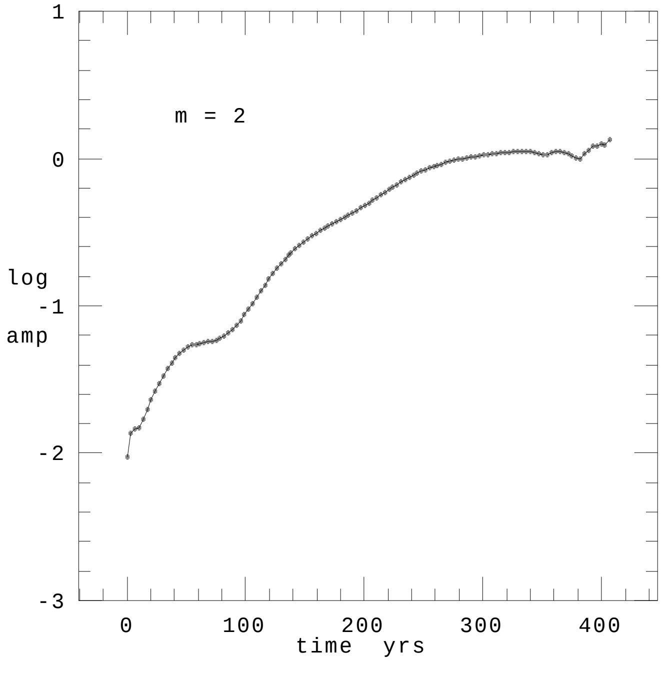

FIG. 6.—Time evolution of the amplitude of the $m = 2$ density mode at a fixed radius of 8.13 AU for model n. Prior to reaching saturation, the amplitude grows roughly exponentially with time.

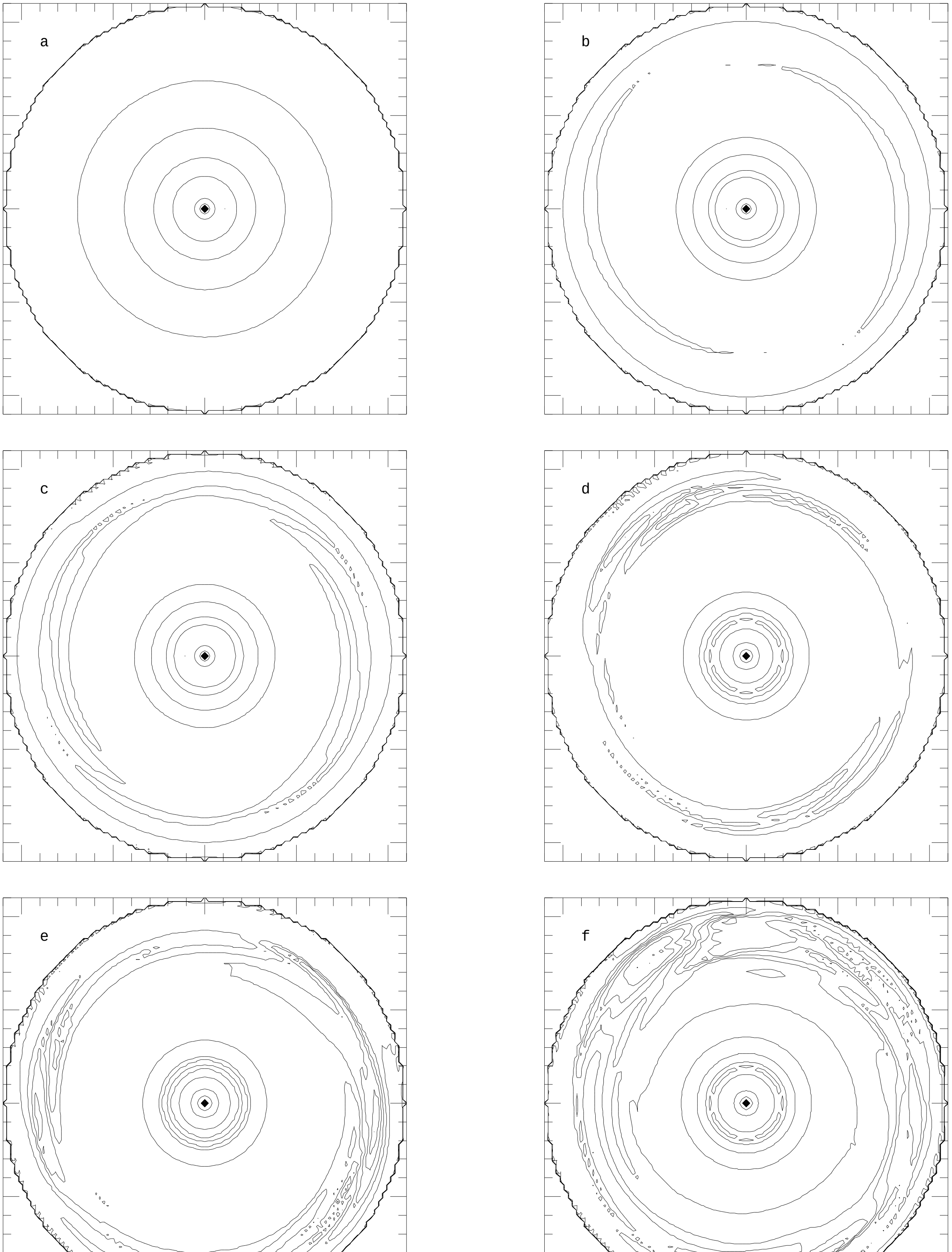

FIG. 7.—Time evolution of the midplane density in model n: (*a*) 0.0 yr, (*b*) 169 yr, (*c*) 255 yr, (*d*) 344 yr, (*e*) 379 yr, and (*f*) 407 yr. The diamond denotes the solar-mass protostar at the center of the 20 AU diameter disk. Each contour represents a factor of 2 change in density, with the maximum density, ranging from 2×10^{-9} g cm^{-3} in (*a*) to 1×10^{-8} g cm^{-3} in (*f*), generally increasing toward the center. The initially nearly axisymmetric, counterclockwise-rotating disk develops trailing spiral arms that continue to grow in amplitude in the outer disk and form two GGPPs, seen at 11 o'clock and 5 o'clock in (*f*).

axisymmetry is sheared into two trailing spiral arms, which continue to grow in amplitude while remaining confined to the outer disk. By 255 yr, the spiral arms have grown strong enough that their gravitational torques are removing orbital angular momentum from inside a radius of 8 AU and depositing it outside 8 AU at a rate of $\sim 10^{40}$ g cm^2 s^{-2}, a rate high enough to transport all the disk's angular momentum ($J_d = 3.0 \times 10^{52}$ g cm^2 s^{-1}) in $\sim 10^5$ yr. As the spiral arms increase in amplitude, their peak gravitational torques continue to grow, lowering the timescale for angular momentum transport to $\sim 10^4$ yr by 344 yr. The spiral arms are quite distinct by this time, with the phase of the maximum density of the two-armed spiral being well ordered (Fig. 8).

As the spiral arms continue to grow in amplitude, their maximum density regions become increasingly prominent and well defined. At this point, the banana-shaped clumps may be properly called GGPPs. For model n, at a time of 344 yr (Fig. 7*d*), the two clumps each have masses of about 2 $M_{\rm Jup}$, where all the gas surrounding each density maximum and with a density of at least $0.1\rho_{\rm max}$ is included in the total. However, thereafter one GGPP begins to grow faster than the other. By 351 yr, one GGPP contains 4 $M_{\rm Jup}$, while other still has only 2 $M_{\rm Jup}$ of gas and dust. By 379 yr, the smaller mass GGPP has begun to lose mass, falling to 1 $M_{\rm Jup}$, while the more massive GGPP attains 6 $M_{\rm Jup}$ by 407 yr. By then the mass ratio is greater than 10:1. One measure of this differential growth is evident in Figure 9: the $m = 1$ mode begins to grow rapidly after the GGPPs become well defined, allowing one GGPP to grow at the expense of the other. Even though the disk was seeded with an $m = 2$ perturbation, the final outcome may well be the formation of a single GGPP moving on an initially circular orbit in the outer disk.

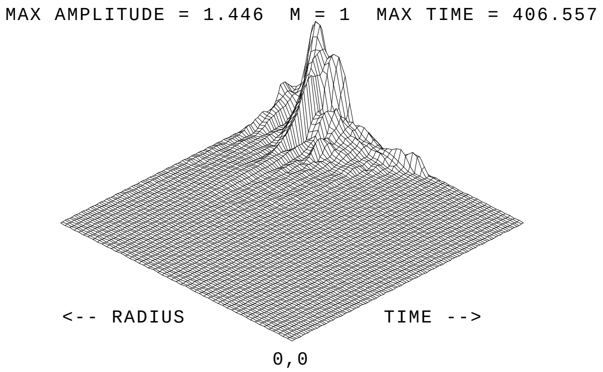

FIG. 9.—Time evolution of the amplitude of the $m = 1$ density mode as a function of disk radius in the unstable disk model n ($Q_{\rm min} = 0.94$), as in Fig. 3. The $m = 1$ amplitude rises rapidly after ~ 300 yr, when two GGPPs begin to form, leading to preferential growth of one GGPP.

The gravitational torques of the GGPPs begin to open a gap in their vicinity, as is evident in model n's surface density profile (Fig. 10). The initially power-law σ profile has developed a peak at the radius of the GGPPs surrounded by troughs and berms. Because of the inviscid nature of the calculations, there is no physical disk viscosity to resist the formation of gaps (cf. Lin & Papaloizou 1986; Ward & Hourigan 1989); evidently, whatever numerical viscosity is present is unable to counteract the growing gravitational torques.

The extent of the gravitational interactions of the GGPPs with the disk is consistent with the predictions of spiral density wave theory (see, e.g., Binney & Tremaine 1987). For an *m*-armed spiral pattern rotating with angular veloc-

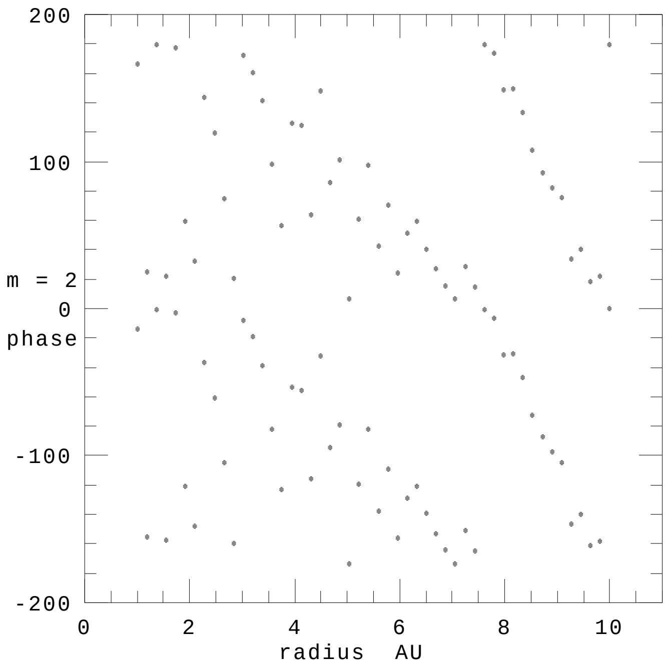

FIG. 8.—Phase (in degrees) of the $m = 2$ mode of the density in the disk midplane at 344 yr for model n. While the mode is completely disordered in the inner disk, a clear two-armed spiral pattern in evident in the outer disk, centered about 8.5 AU.

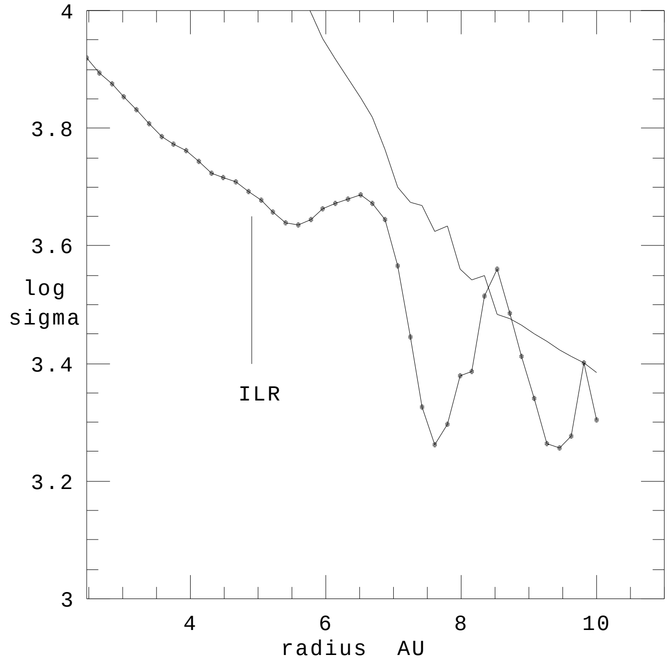

FIG. 10.—Log of the surface density (σ, in g cm^{-2}) as a function of radius for model n at 379 yr. Solid line is minimum surface density needed to have the Toomre $Q < 1$. The GGPPs orbit around 8.5 AU, at the local peak in σ. The inner Lindblad resonance at 4.9 AU marks the inner boundary of the region disturbed by the GGPPs.

ity Ω_b, the frequency of intersections per epicycle of a fluid element in the disk is

$$\nu = \frac{m(\Omega - \Omega_b)}{\pm\kappa} .$$

The Lindblad resonances occur when $\nu = 1$, yielding

$$\kappa = \pm m(\Omega - \Omega_b) .$$

For a Keplerian disk, the epicyclic frequency $\kappa = \pm\Omega$, the disk angular velocity, so

$$\Omega = \frac{m}{m \pm 1}\,\Omega_b .$$

Density waves can propagate between the outer and inner Lindblad resonances. For a two-armed spiral ($m = 2$), the outer and inner Lindblad resonance (OLR and ILR, respectively) fall at

$$\Omega = \tfrac{2}{3}\Omega_b \quad \text{and} \quad 2\Omega_b .$$

For model n at 379 yr, $\Omega_b = 10^{-8}$ rad s^{-1}, yielding an OLR beyond 10 AU, outside the computational volume, and an ILR at 4.9 AU. As shown in Figure 10, the ILR marks the approximate inner boundary of the portion of the disk where the gravitational torques of the GGPPs have disturbed the initial σ profile. Figure 11 shows that the phase of the $m = 2$ mode, which demonstrates nearly uniform rotation at the location of the GGPPs around 8.5 AU, traces a two-armed spiral pattern inward to the ILR, where the phase becomes disordered.

4.4. *Varied Local Thermodynamics*

As shown by Boss (1997b), the disk instability can proceed for $\gamma = 7/5$ as well as for $\gamma = 1$, the two extremes for adiabatic or isothermal behavior, respectively. While it would be interesting to explore the effects of disk heating and cooling processes during the instability, the fact that the instability proceeds regardless of γ implies that the instability is a robust one, provided that $Q_{\min}$ is low enough.

Figure 12 depicts the growth of the $m = 2$ mode in model z, with $\gamma = 7/5$ and $Q_{\min} = 0.96$, a value close to that of the $\gamma = 1$ model n ($Q_{\min} = 0.94$). The evolution of this unstable disk model is very similar to that of model n (Fig. 5), except for taking a somewhat longer time to develop to the same maximum amplitude (about 430 yr for model z, compared to 361 yr for model n). The maximum temperatures inside the GGPPs exceed 300 K by 702 yr, well above the initial outer disk temperature of 75 K. Evidently adiabatic heating during the instability does not prevent GGPP formation. By 702 yr, the masses of the GGPPs have risen to $\sim 8\,M_{\rm Jup}$ apiece, and the average temperature inside the GGPPs is 250 K.

Figure 13 shows the stability results as a function of $Q_{\min}$ and γ for initial $m = 2$ density perturbations. While the models only coarsely sample this parameter space, there is no evidence that critical values of $Q_{\min}$ depend on γ, i.e., it appears that $Q_{\min} < 1$ is sufficient for an unstable disk, independent of γ. Similarly, it appears that the critical value of $Q_{\min}$ for marginal stability, $Q_{\min} < 1.5$, may also be approximately independent of γ. Because several marginal disks were evolved far enough in time for GGPP formation to occur (models o and c), it thus appears that a sufficient condition for GGPP formation is $Q_{\min} < 1.5$. The critical value of $Q_{\min} < 1$ for disk instability is formally identical to that advanced by Toomre (1964) for a stellar disk, though Toomre's analysis involved axisymmetric disturbances rather than $m = 2$ nonaxisymmetric perturbations, as would be appropriate for very tightly wound spiral density waves.

4.5. *Varied Density Perturbations and Protostellar Wobbles*

The previous models (Table 1) were all started with a dominant $m = 2$ density perturbation of amplitude $a_2 = 0.01$, initially in phase throughout the disk, which predisposes unstable disks to forming pairs of GGPPs. Protoplanetary disks need not be similarly seeded, so it is important to learn what happens when disks are given different initial density perturbations.

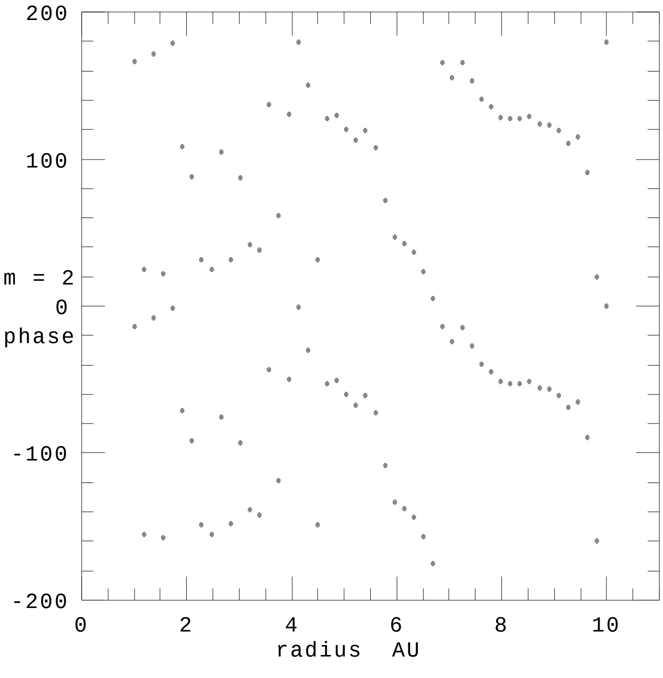

FIG. 11.—Phase (in degrees) of the $m = 2$ mode of the density in the disk midplane at 406 yr for model n. The mode rotates nearly uniformly in the vicinity of the GGPPs at 8.5 AU and maintains a spiral pattern inward to the ILR at 4.9 AU, inside of which the pattern becomes disordered.

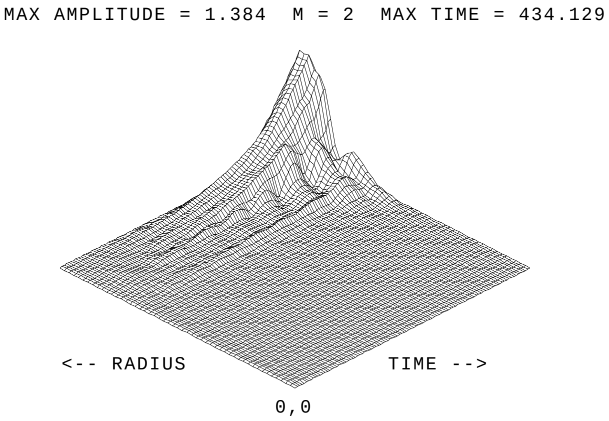

FIG. 12.—Time evolution of the amplitude of the $m = 2$ density mode as a function of disk radius in the $\gamma = 7/5$ disk model z ($Q_{\min} = 0.96$), as in Fig. 3. Unstable disks occur even for adiabatic local thermodynamics, provided that $Q_{\min} < 1$.

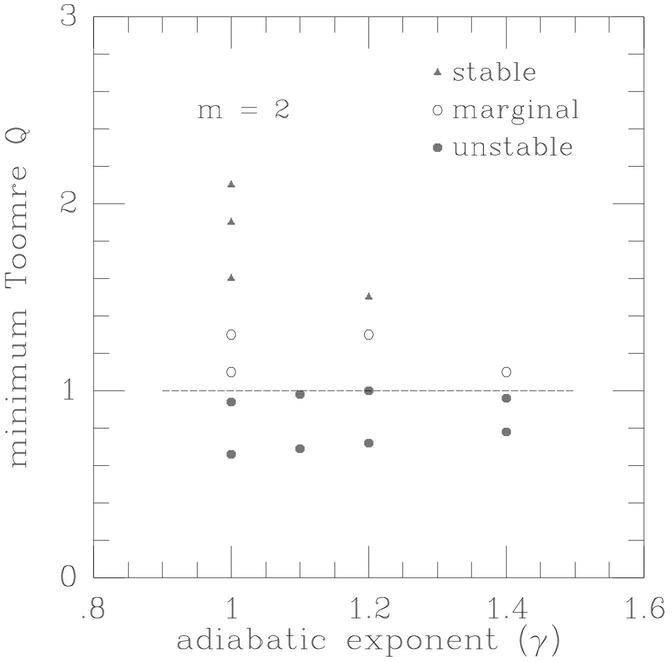

FIG. 13.—Disk stability to $m = 2$ density perturbations as a function of the assumed thermodynamics (γ) and the minimum value of the disk's Toomre Q parameter. Unstable disk models all have $Q < 1$ (*dashed line*), consistent with Toomre (1964). However, marginally unstable disks seem to be able to form GGPPs for $1 < Q < 1.5$.

Several variations of model o were calculated in order to learn the behavior of low-order $m \neq 2$ modes in a marginally stable disk. Model a was identical to model o except for having an initial $m = 1$ perturbation instead of $m = 2$. The $m = 1$ mode began to grow immediately, indicating instability, because the $m = 1$ mode has a longer wavelength in the azimuthal direction than the $m = 2$ mode and so effectively can involve more self-gravitating gas than the marginally stable $m = 2$ mode (Boss 1982). The $m = 1$ mode continued to grow strongly and by 650 yr had formed a single GGPP moving in a circular orbit at 9 AU with a mass of $\sim 10\ M_{\rm Jup}$.

Wobbling protostars preferentially excite the $m = 1$ mode (Adams et al. 1989). Model aw was identical to model a, except that the protostar was allowed to wobble back and forth in order to preserve the location of the system's center of mass precisely at the coordinate origin. While the outer disk evolved in much the same manner as in model a, forming one GGPP at 9 AU, the inner disk behaved considerably differently. The wobbling protostar served as a feedback mechanism for the growth of $m = 1$ nonaxisymmetry in the inner disk: the greater the wobble of the protostar, the greater the differential gravitational force on the disk, leading to larger disk asymmetry and hence a larger protostellar wobble. The protostellar wobble in model a led to the formation of a second GGPP at an orbital radius of ~ 5 AU, a suggestive location considering the orbital radius of Jupiter.

Model ah began with an initial $m = 16$ density perturbation and behaved very similarly to model az, which had no density perturbation at all. Both models ah and az appeared to be stable, with the amplitudes of the low-m modes ($m = 1$, 2, 3) remaining quite small (< 0.005). The short wavelength of the initial $m = 16$ perturbation in model ah accounts for its stability, consistent with the marginal stability of the $m = 2$ mode compared to the $m = 1$ mode. It is noteworthy that the low-level modes $m = 1$, 2, 3, ... initially present as noise in model ah did not grow either, even in the presence of the wobbling protostar. Evidently instability in these models requires a density perturbation with a finite amplitude (~ 0.01) that is initially in phase throughout the disk. Including a wobbling protostar was insufficient to lead to growth of nonaxisymmetry. The stability of models ah and az (calculated to 544 yr and 311 yr, respectively) is also strong evidence of the stability of these disk models to numerical noise, and hence of the absence of artificial fragmentation caused by the growth of unphysical effects (Truelove et al. 1997).

Finally, model aa was started with perhaps the most realistic initial conditions—all four low-order modes ($m = 1$, 2, 3, 4) were given an initial amplitude of 0.01, and the protostellar wobble was included. Model aa was unstable to the growth of the $m = 1$ mode, in both the middle and outer disk (Fig. 14); the growth of the $m = 1$ mode in the outer disk was very similar to that of model a, implying independent growth of the modes in the middle and outer disk. The disk eventually formed two GGPPs, one around 6 AU and one at 10 AU (Fig. 15), with masses of $\sim 18\ M_{\rm Jup}$ and $\sim 10\ M_{\rm Jup}$, respectively. Considering that $\sim 130\ M_{\rm Jup}$ of disk gas is available for GGPP formation in these models, the rather large masses of the GGPPs in model aa are perhaps not too surprising.

4.6. *Varied Azimuthal Resolution*

In order to study the dependence of the growth of noise on the spatial resolution of the models, several models were calculated with variations in the number of azimuthal grid points, N_ϕ. Three models were calculated, each of which was identical to the initially axisymmetric model az (with $N_\phi = 64$), except for having $N_\phi = 8$, 16, and 32, respectively. After about 300 yr, the amplitude of the $m = 2$ mode in the usually noisy inner disk had grown to the range 0.1–1 in each of the lower resolution models, whereas in model az it had grown to only 0.01 and then decreased again. Evidently, poor spatial resolution can exacerbate the growth of

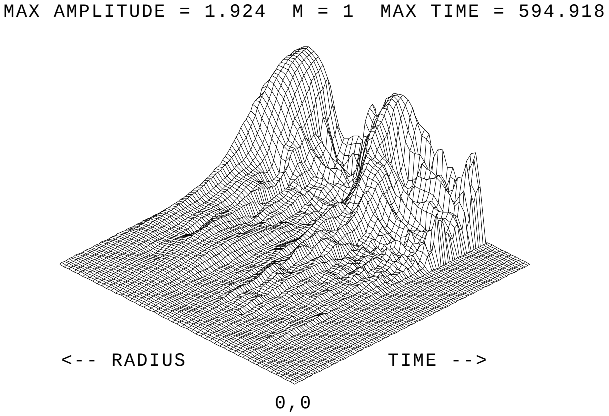

FIG. 14.—Time evolution of the amplitude of the $m = 1$ density mode as a function of disk radius in model aa ($Q_{\rm min} = 1.1$), as in Fig. 3. The inclusion of protostellar wobble excites the growth of the $m = 1$ density mode and leads to the formation of a GGPP in the middle of the disk at ~ 6 AU. An additional GGPP forms near the edge of the disk. Numerical noise is evident just outside the boundary of the artificially axisymmetric innermost disk, located at about 2.7 AU in this model.

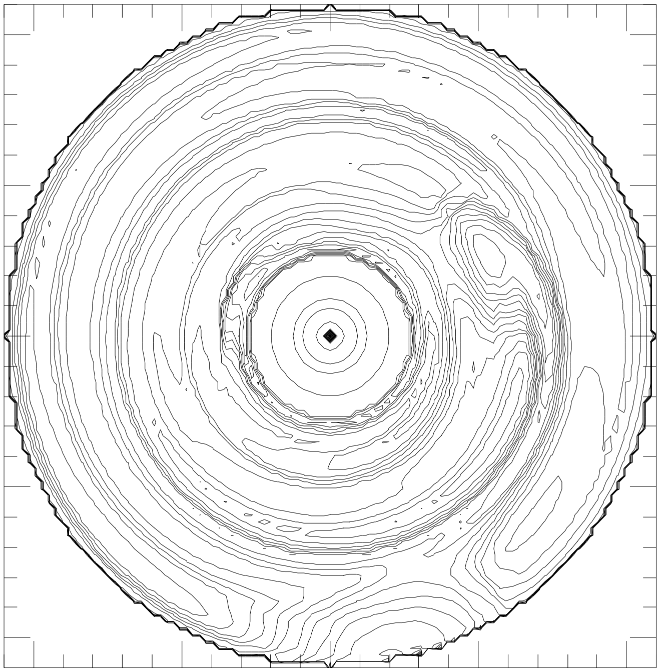

FIG. 15.—Midplane density of model aa at 619 yr, as in Fig. 7. Maximum density is 5×10^{-8} g cm^{-3}. GGPPs are evident at 2 o'clock at 6 AU radius and near 6 o'clock at 9.5 AU radius. Each GGPP has accreted most of the gas initially within its annular ring.

numerical noise in these disks, but with the standard resolution, $N_\phi = 64$, the level of noise is not worrisome.

A model was also computed with double the standard azimuthal resolution. Model n2 had effectively $N_\phi = 128$ (achieved by employing symmetry through the rotation axis, restricting the model to even *m* modes) but was otherwise identical to model n, with $N_\phi = 64$ and no symmetry assumptions. Model n2 was intended to address the issue of artificial fragmentation raised by Truelove et al. (1997) in the context of isothermal cloud collapse. Truelove et al. (1997) suggested that artificial fragmentation could occur whenever the cell size Δx exceeds $\frac{1}{4}$ of the local Jeans length λ_J, where

$$\lambda_J = \left(\frac{\pi c_s^2}{G\rho}\right)^{1/2} .$$

For a spherical coordinate grid, there are three different grid spacings, depending on the coordinate direction:

$$\Delta x_r = \Delta r , \quad \Delta x_\theta = r\Delta\theta ,$$

and

$$\Delta x_\phi = r \sin\theta \, \Delta\phi .$$

For a spherical coordinate grid, a single grid spacing (comparable to that employed by Truelove et al. 1997) can be defined as

$$\Delta x = (\Delta x_r \, \Delta x_\theta \, \Delta x_\phi)^{1/3} ,$$

where Δx is the width of a cubical cell with approximately the same volume as the spherical coordinate cell. If the Jeans *mass* is the physically relevant quantity rather than the Jeans *length*, then Δx should be the applicable cell size for the present models. In the adaptive mesh refinement models of Truelove et al. (1997), their Cartesian grid is locally uniform, so the Jeans mass criterion is identical to the Jeans length criterion. While it could be argued that the Jeans mass is the more appropriate criterion for any nonuniform grid, a more conservative approach is to try to satisfy both criteria.

The initial values of these various lengths are listed in Table 3 for both models n and n2, at the GGPP-forming radius of 8 AU in the midplane. Table 3 shows that for model n, Δx_r, Δx_θ, and Δx are each considerably less than $\lambda_J/4$, but $\Delta x_\phi \approx \lambda_J/4$. If Δx is the proper length, then model n satisfies the Truelove et al. (1997) criterion, but if Δx_ϕ is the proper length, then the criterion is only marginally satisfied. However, for model 2n, all four grid spacings are comfortably less than $\lambda_J/4$.

If artificial fragmentation is responsible for GGPP formation in model n, then model n2 should not undergo GGPP formation. Instead, model n2 behaves very similarly to model n, as shown in Figure 16. The amplitude of the $m = 2$ mode in model n2 grows very nearly identically to the $m = 2$ mode in model n (Fig. 6). Figures 6 and 16 strongly imply that the hydrodynamic solutions for the $m = 2$ mode have converged, and that the spatial resolution is quite adequate for these exploratory models of the initial

TABLE 3

INITIAL JEANS LENGTH AND GRID SPACINGS FOR TWO MODELS IN THE DISK MIDPLANE AT AN ORBITAL RADIUS OF 8 AU

Parameter	Model n (cm)	Model n2 (cm)
λ_J	4.9×10^{13}	4.9×10^{13}
$\lambda_J/4$	1.2×10^{13}	1.2×10^{13}
Δx_r........	2.8×10^{12}	2.8×10^{12}
Δx_θ	7.1×10^{11}	7.1×10^{11}
Δx_ϕ	1.2×10^{13}	6.0×10^{12}
Δx	2.9×10^{12}	2.3×10^{12}

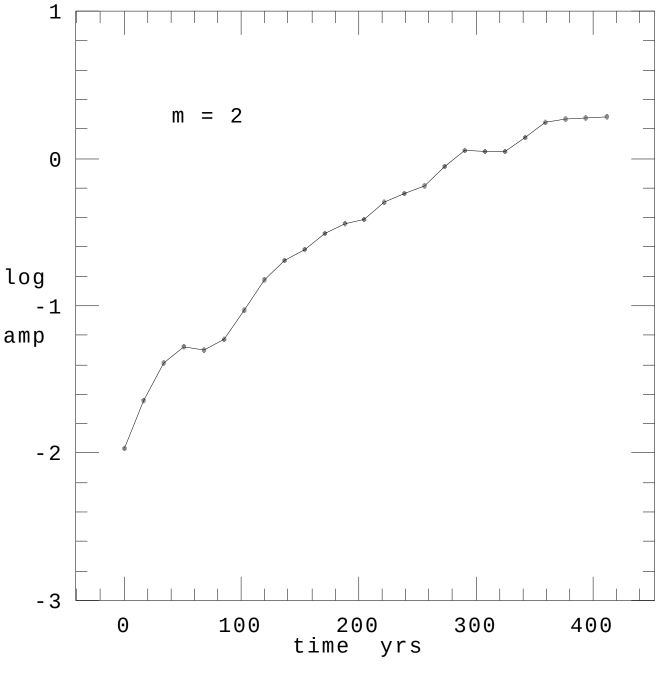

FIG. 16.—Time evolution of the amplitude of the $m = 2$ density mode at a fixed radius of 8.13 AU for model n2, with twice the azimuthal resolution of model n (cf. Fig. 6).

phases of GGPP instability. Following the subsequent evolution of the GGPPs, however, will require considerably greater spatial resolution of the GGPPs.

5. DISCUSSION

5.1. *Comparison with Previous Work*

Previous numerical studies of circumstellar disks have shown that marginally unstable disks can evolve through the growth of trailing spiral arms and the subsequent transport of mass and angular momentum caused by the gravitational torques associated with the spiral arms (Cassen et al. 1981; Boss 1984; Tomley et al. 1991; Adams & Benz 1992; Laughlin & Bodenheimer 1994; Tomley, Steiman-Cameron, & Cassen 1994; Pickett, Durisen, & Davis 1996; Laughlin & Różyczka 1996; Drimmel 1996; Pickett, Durisen, & Link 1997; Laughlin, Korchagin, & Adams 1997; Nelson et al. 1998). Here we compare the present results to these previous studies.

Cassen et al. (1981) used an N-body code to model two-dimensional, infinitely thin, isothermal disks with $\sigma_i \propto r^{-1}$. They found that disk stability depended strongly on the ratio of the disk mass to the central star mass. For a 50 AU radius, $M_d = 0.1\ M_\odot$ disk orbiting a solar-mass star, a situation somewhat comparable to that of the present models, a model with a disk temperature of 100 K and $Q_{\rm min} = 3.7$ was stable, consistent with the present results. Tomley et al. (1991, 1994) continued these N-body calculations but considered only disks with small radii ($R < 1.8$ AU) that cannot be compared to the present models, in which the innermost disk remained axisymmetric because of the high inner disk temperatures (Fig. 1).

Laughlin & Bodenheimer (1994) studied the three-dimensional evolution of a disk with a mass ($\approx 0.5\ M_\odot$) comparable to that of its central protostar, using a smoothed particle hydrodynamics (SPH) code and assuming locally isothermal thermodynamics, as in the present $\gamma = 1$ models. They found that nonaxisymmetry grew in a model with $Q_{\rm min} = 1.3$, forming trailing spiral arms throughout the evolution, which lasted for over 1 rotation period at the initial edge of the disk (230 AU). Another model with $Q_{\rm min} \approx 1$, however, fragmented into two blobs in 0.1 the elapsed time of the $Q_{\rm min} = 1.3$ model. While these results apply to much larger radii disks than the present models, they are fairly consistent in terms of their $Q_{\rm min}$ dependence.

Laughlin & Różyczka (1996) used a two-dimensional code to systematically study massive, thin disks with $\gamma = 2$ and 5/3 thermodynamics, finding that $Q_{\rm min} \approx 1.15$ separated disks that fragment from those that evolve through spiral arms. The main difference with the present models appears to be the eventual fragmentation of at least two of the marginally unstable disks (models o and c) in the present study, compared to the persistent spiral structure in Laughlin & Różyczka's (1996) $Q_{\rm min} > 1.15$ models. The exact reason for this discrepancy is unclear, but is presumably related to the differences in the calculations. For example, models o and c were evolved to a time greater than $20P_{10}$, whereas Laughlin & Różyczka's (1996) models generally were evolved for about 10 outer disk rotation periods. Other possibly important differences include the values of γ employed, the dimensionality of the calculations, the initial disk σ profiles (Gaussian in the case of Laughlin & Różyczka), and hence the initial thermal profiles.

Drimmel (1996) performed two-dimensional SPH studies of massive, thin disks with $\gamma = 2$ and 5/3 and $Q_{\rm min} = 1.15$, finding that when the disk mass equalled the central mass, growth of the $m = 1$ mode could produce self-gravitating clumps. Clump formation was inhibited, however, when the disk mass was only one-third that of the central object. The initial surface density was $\sigma \propto e^{-r/r_s}$, and the calculations explicitly included the effects of viscosity, so once again the models are not directly comparable to the present work.

Pickett et al. (1996, 1997) computed three-dimensional models of polytropic ($\gamma = 5/3$) disks with varied angular momentum distributions. These disks included central concentrations similar to central protostars, and when scaled to a total mass of $\sim 1\ M_\odot$, the disks had radii of up to a few AU. Models with $Q_{\rm min} \sim 1$ were unstable to the formation of nonaxisymmetric structure, though it was unclear if the system would breakup into several objects.

Continuing the effort initiated by Adams & Benz (1992), Nelson et al. (1998) presented two-dimensional models of thin, locally isothermal disks with radii of 50 AU. As in the present study, Nelson et al. (1998) found that disks with $Q_{\rm min} < 1.5$ broke up into clumps. However, some of the SPH disk models broke up into as many as 33 clumps, considerably larger than the number of GGPPs found in the present unstable models. Because of the differences in the initial disk models, a more precise comparison is not possible, but the results of Nelson et al. (1998) regarding clump formation for $Q_{\rm min} < 1.5$ and relative stability of disks for $Q_{\rm min} > 2$ are qualitatively consistent with the present models.

5.2. *Triggering the GGPP Instability*

In general, the previous numerical studies did not encounter clump formation unless the disk was fairly unstable, a situation that is arguably unobtainable, because marginally unstable disks are believed to evolve toward a more stable configuration (e.g., by transporting disk mass inward onto the stabilizing protostar). As a circumstellar disk builds up by the infall of molecular cloud gas and dust, then, the disk is expected to transport mass inward and remain marginally unstable. However, the present finding that the marginally unstable disk models o and c eventually began to break up into GGPPs casts some doubt on whether or not all marginally unstable disks behave in the same manner. If marginally unstable disks can evolve in such a way as to form GGPPs rather than only transport mass inward, then GGPP formation may be a natural outcome of the evolution of protostellar and protoplanetary disks.

On the other hand, if $Q_{\rm min}$ has to be considerably decreased in order to transform a spiral-arm–evolving disk into a GGPP-forming disk, it could then be argued that the GGPP instability is unlikely to ever occur, at least in the absence of some mechanism that could suddenly (faster than the disk can evolve by spiral density waves) push the disk over into the GGPP formation regime. One such mechanism could be episodic accretion of mass by a marginally unstable disk from the infalling remnant of the parent molecular cloud core.

Boss (1997b) suggested that clumpy accretion of infalling envelope gas could drive $Q_{\rm min}$ below 1. Direct observations of infalling clumps in the B335 molecular cloud by Velusamy et al. (1995) imply that several $M_{\rm Jup}$ of gas could be added to the disk in a single clump over a time period of less than 10^3 yr. Spiral density waves presumably would be trying to transport away this excess mass at the same time that the GGPPs are trying to form, and the outcome of this competition remains uncertain in the absence of a detailed calculation. If enough mass is dumped more or less instantaneously onto a marginally unstable disk, however, the GGPP instability should be triggered.

On the basis of the evidence for short-lived radioactivities in primitive meteorites, it has been suggested that the collapse of the presolar cloud was triggered by the arrival of a supernova shock wave carrying freshly synthesized radioactive nuclei (Cameron & Truran 1977). Numerical models have shown that shock-wave triggering accelerates the collapse of the cloud envelope onto the protostar and its disk (Boss 1995). The implications of shock-wave triggering for GGPP formation remain to be explored, however.

5.3. *Stability of GGPPs*

The present models are necessarily limited to the earliest phase of GGPP formation in unstable disks. The subsequent contraction and evolution of the GGPPs is therefore uncertain, and in particular whether or not giant planets can be formed from the GGPPs is unproven. At a minimum, though, the GGPPs must be self-gravitating and tidally stable with respect to the protostar's gravitational forces if giant planets are to result.

Whether or not the GGPPs are gravitationally bound can be determined by comparing their masses with the relevant Jeans mass for their densities and temperatures. If the Jeans mass M_J is defined as the mass of a uniform density sphere with a diameter equal to the Jeans length, then in cgs units,

$$M_J = 1.3 \times 10^{23} \left(\frac{T}{\mu}\right)^{3/2} \rho^{-1/2} ,$$

where $\mu \approx 2.3$ is the mean molecular weight. For a typical GGPP with a temperature of ~ 100 K and a density of $\sim 10^{-8}$ g cm^{-3}, $M_J \approx 0.2\ M_{\rm Jup}$. Considering that the GGPPs in the present models typically have masses of several $M_{\rm Jup}$ or more, it is clear that these clumps are gravitationally bound, even when their internal temperatures are ~ 300 K.

The free-fall time $t_{\rm ff}$ for a uniform density sphere is defined as

$$t_{\rm ff} = \left(\frac{3\pi}{32 G \rho}\right)^{1/2} .$$

For a GGPP with a density of 10^{-8} g cm^{-3}, $t_{\rm ff} \sim 0.7$ yr, considerably shorter than the orbital period. Hence the GGPPs can be expected to begin to contract to higher densities on a short timescale.

Besides internal gas pressure and centrifugal effects, GGPP contraction will be resisted by the tidal forces exerted by the protostar and the disk. A tidal stability criterion was derived by DeCampli & Cameron (1979) for a protoplanet of mass M_p orbiting a protostar of mass M_s, inside a disk with a surface density distribution that results in constant orbital velocity, rather than the $v_\phi \propto r^{-1/2}$ dependence that characterizes a Keplerian disk. Because the disk mass is much smaller than the protostar mass in the present models, the disk mass can be neglected in the DeCampli & Cameron (1979) derivation. This approximation yields a critical tidal radius r_t (the inner Lagrange point) that must exceed the planetary radius R_p in order for the planet to be stable with respect to the tidal force of the protostar:

$$R_p < r_t = r_p \left(\frac{M_p}{3M_s}\right)^{1/3} ,$$

where r_p is the orbital radius of the planet. For a solar-mass protostar, tidal stability then requires

$$R_p < 0.069 r_p \left(\frac{M_p}{M_{\rm Jup}}\right)^{1/3} .$$

For the $\sim 6\ M_{\rm Jup}$ GGPP located at ~ 8.5 AU in Figure 7*f* (model n) to be tidally stable, its radius must be less than ~ 1.1 AU. While the extent of the GGPP in Figure 7*f* is perhaps 1 or 2 AU in the azimuthal direction, in the tidally critical radial direction its radius is just less than 1 AU. Hence the bulk of the gas is indeed gravitationally bound to the GGPP. Similarly, for the $\sim 18\ M_{\rm Jup}$ GGPP located at ~ 6 AU in Figure 15 (model aa) to be tidally stable, its radius must not exceed ~ 1.1 AU, which is also the case.

To the extent that a portion of a GGPP does extend beyond the tidal radius, this outlying material would not participate in the contraction toward giant planet densities, reducing the maximum mass of the giant planet. Considering the rather large masses obtained in the present models ($\sim 10\ M_{\rm Jup}$), this does not appear to be a major concern, and in fact could be viewed positively, as a means of reducing somewhat the estimated masses of the final planets to a range perhaps closer to that of the observed solar system and extrasolar giant planets.

5.4. *Core Formation in a GGPP*

Slattery, DeCampli, & Cameron (1980) showed that liquid drops of iron or rock would sink to form a core in a GGPP. Stevenson (1982) agreed that the metals initially present would rain out to form a core but noted that the resulting core mass would be considerably smaller than the estimates for core masses prevalent at the time. Stevenson (1982) pointed out that if one tried to boost the core mass through impacts of small bodies after the planet had already formed, the *late-added* iron and silicate would be soluble in the planet's adiabatic envelope and would not rain out to the core. Coupled with the failure of the GGPP model to explain the similarity in masses of the cores of the giant and outer planets, Stevenson (1982) justifiably concluded that the evidence favored the core accretion mechanism.

A 1 $M_{\rm Jup}$ GGPP of solar composition ($Z = 0.02$) contains $\sim 6\ M_\oplus$ of metals, while a Saturn-mass GGPP would contain $\sim 2\ M_\oplus$ of metals. If the dust grains and the gaseous metal component manage to find their way to the center of the GGPPs, then ice/rock cores of sufficient mass to match the latest interior models (implying a 0–12 $M_\oplus$ core for Jupiter and a 1–13 $M_\oplus$ core for Saturn) could form. Sedimentation of the dust in a solar-composition GGPP might then be sufficient to explain the core masses of Jupiter and Saturn, removing one of the primary objections to the GGPP mechanism (Boss 1997b). Considering the present uncertainties in giant planet core masses, it is unclear if the core masses of the giant planets truly are similar to each other or to those of the outer planets, possibly removing the second of the objections of Stevenson (1982).

Following the formation phase depicted in the present calculations, the GGPPs will slowly contract toward planetary densities. Bodenheimer et al. (1980) calculated the evolution of isolated GGPPs in spherical symmetry (one-dimensional) and found that starting from a central temperature of $T_c = 100$ K and a central density of 10^{-9} g cm^{-3} (similar to the present models), a 1 $M_{\rm Jup}$ GGPP would contract for $\sim 10^5$ yr before reaching $T_c = 1000$ K. During this period, conditions should be hospitable for the survival of iron and silicate dust grains and for the continued insolubility of water in molecular hydrogen (Stevenson & Fishbein 1981), the dominant gas in a GGPP.

While the gaseous component of the GGPP is contracting on a timescale of 10^5 yr, dust grains can grow by collisional coagulation and settle to the center of a nonturbulent GGPP in much the same way that dust grains are envisioned to settle toward the midplane of the solar nebula (see, e.g., Weidenschilling 1988). Small dust grains (with radius $a < 1$ m) moving through a gas of density ρ obey the Epstein drag law, yielding a response time t_e to gas drag

$$t_e = \frac{a\rho_a}{\rho c_s} ,$$

where ρ_a is the density of the dust grain itself. The terminal settling velocity V of a dust grain subject to a gravitational acceleration a_g is $V = a_g t_e$. For a uniform density, spherical GGPP of density ρ_p, the gravitational acceleration at a radius R inside the GGPP is

$$a_g = -\tfrac{4}{3}\pi G \rho_p R .$$

The terminal velocity is then

$$V = a_g t_e = -\frac{4\pi a G \rho_a R}{3 c_s} ,$$

leading to a time for reaching the center of the GGPP of

$$\tau_c = \frac{R}{|V|} = \frac{3 c_s}{4\pi a G \rho_a} .$$

Because the terminal velocity depends on size, larger grains will settle faster than smaller grains. This differential settling will cause the larger grains to sweep up the smaller grains, leading to growth of the grain mass m_a at the rate

$$\frac{dm_a}{dt} = \pi a^2 \rho_d |V| ,$$

where ρ_d is the space density of dust grains, taken to be some fraction f of the gas density ($\rho_d = f\rho_p$). The characteristic growth time $\tau_a = m_a/(dm_a/dt)$ for a dust grain is

$$\tau_a = \frac{c_s}{\pi f \rho_p G R} .$$

For $R = 0.5$ AU (initially in the middle of the GGPP), $T = 100$ K, $f = 0.01$, and $\rho_p = 10^{-8}$ g cm^{-3}, the timescale for dust grain growth in mass by a factor of e is about 10 yr. The maximum size a dust grain can grow to before reaching the center of the GGPP is

$$a_{\rm max} = \frac{f \rho_p R_{\rm max}}{4\rho_a} ,$$

yielding $a_{\rm max} \sim 1$ m for $R_{\rm max} = 1$ AU and $\rho_a = 3$ g cm^{-3}. The time to grow to size a starting from size a_o is $\sim 3\tau_a \ln(a/a_o)$. For $a_0 = 0.1$ μm, centimeter-sized particles will form within $\sim 10^3$ yr. Centimeter-sized particles can then settle to the center of the GGPP in a time $\tau_c \sim 10^3$ yr. Because both of these timescales are considerably shorter than the $\sim 10^5$ yr needed for the gaseous envelope to contract, it appears to be quite likely that GGPPs will form solid cores (Boss 1997b).

In addition to the settling of dust grains initially contained within a GGPP, a growing GGPP can acquire further solids by intercepting the radially inward flux of particles that have coagulated outside its orbit. Cameron (1995) showed that centimeter-sized particles would spiral inward past radii of 5 AU in $\sim 10^5$ yr, while meter-sized bodies would be intercepted in $\sim 10^3$ yr. Radially inward orbital migration, caused by gas drag, could then feed additional material to a growing GGPP core, prior to the GGPP reaching temperatures high enough to prevent sedimentation of these solids to the core.

While the present models do not directly address this issue, it should be emphasized that Uranus and Neptune probably did not form by the GGPP instability mechanism (cf. Cameron 1978). The falling surface density of the disk may lead to insufficient mass much beyond 10 AU for the GGPP instability to operate, even at the inferred minimum outer disk temperatures of 25–50 K (Mumma 1996). Uranus and Neptune would then have to have formed well after Jupiter and Saturn, and well after the residual solar nebula had been removed, by the slow process of collisional accumulation, as in the inner solar system.

5.5. *Testing the GGPP Hypothesis*

Finally, it is important to note that a definitive test for deciding between the core accretion and GGPP instability mechanisms for giant planet formation is possible. As it is presently understood (see, e.g., Pollack et al. 1996), the core accretion mechanism requires at least $\sim 10^6$ yr to form 10 $M_\oplus$ cores, and up to $\sim 10^7$ yr for a core to accrete a massive gaseous envelope. Hence if giant planets form by core accretion, they should not reach Jupiter mass until their stars are $\sim 10^7$ yr old. The GGPP instability, on the other hand, is most likely to occur during the protostellar or very early pre–main-sequence phase of evolution, i.e., at ages of $\sim 10^5$ yr. Nearby star-forming regions contain young stars with ages in the range $\sim 10^5$–10^7 yr. A search for signs of Jupiter-mass companions to these young stars might then define the epoch of giant planet formation and so determine the mechanism of their creation (Boss 1998b).

6. CONCLUSIONS

The results of this paper may be summarized in the context of a list of the apparent advantages and disadvantages of the GGPP mechanism for giant planet formation.

6.1. *Advantages*

The GGPP instability occurs quickly, within 10^3 yr, with core formation occurring on a similar timescale, and with envelope contraction to planetary densities occurring in $\sim 10^5$ yr. Giant planet formation could thus occur well before the disappearance of the protoplanetary disk.

The mechanism appears to be capable of making fairly massive planets. While the present models do not predict the final giant planet masses, the production of ~ 10 $M_{\rm Jup}$ GGPPs suggests that multiple-$M_{\rm Jup}$ giant planets will result. Lower mass disks should produce lower mass GGPPs, though this remains to be studied.

Ice and rock cores should be able to form inside Jupiter- and Saturn-mass GGPPs with core masses comparable to those inferred in recent models of giant planet interiors.

If GGPP formation is a natural occurrence in marginally unstable disks, then a means for triggering the instability may not be required.

Rapidly falling outer disk surface densities would mean that GGPP formation could occur only in the "middle" of the disk, implying that both the terrestrial and icy outer planets would have to form by the slow process of collisional accumulation.

For the initial disk modeled here, GGPPs form in circular orbits with radii (~ 5 and ~ 10 AU) similar to those of Jupiter and Saturn.

Because GGPPs would form early in the evolution of young stars, searches for Jupiter-mass companions to pre–main-sequence stars should be able to either confirm or reject the GGPP formation mechanism.

6.2. *Disadvantages*

The present models assume the existence of a relatively high-mass disk, with 0.13 $M_\odot$ within 10 AU. Evidence for such massive disks is theoretical at present rather than observational. The present calculations need to be repeated for lower mass disks.

If GGPP formation does not occur as a natural outcome of the evolution of marginally unstable disks, then the instability will require the existence of a suitable trigger, such as nonsteady infall of gas onto the disk.

Giant planets formed by the GGPP mechanism will be roughly solar in composition, unless their gas-rich outer layers are partially removed after core formation occurs, which would then produce metal-rich planets, or unless inward radial migration of solids leads to significant metal enrichment. Impacts of small bodies (e.g., comet Shoemaker-Levy 9) might produce metal-rich atmospheres if mixing is prohibited by radiative zones in the giant planet.

If the cores of the giant and outer planets turn out to be nearly identical, in spite of the trends of recent interior models, the GGPP mechanism would offer no easy explanation.

Undoubtedly other advantages and disadvantages of the GGPP instability will become apparent as this mechanism is further investigated.

I thank Willy Benz, Geoff Bryden, A. G. W. Cameron, Pat Cassen, John Larwood, Jack Lissauer, Lee Mundy, Andy Nelson, Stan Peale, Brian Pickett, and George Wetherill for comments and discussions. The calculations were performed on the DEC Alpha workstations of the Carnegie Institution of Washington. This research was supported in part by the Planetary Geology and Geophysics Program of the National Aeronautics and Space Administration under grant NAG 5-3873.

REFERENCES

Adams, F. C., & Benz, W. 1992, in ASP Conf. Proc. 32, Complementary Approaches to Double and Multiple Star Research, ed. H. A. McAlister & W. I. Hartkopf (San Francisco: ASP), 185
Adams, F. C., Ruden, S. P., & Shu, F. H. 1989, ApJ, 347, 959
Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Güsten, R. 1990, AJ, 99, 924
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Bodenheimer, P., Grossman, A. S., DeCampli, W. M., Marcy, G., & Pollack, J. B. 1980, Icarus, 41, 293
Boss, A. P. 1982, ApJ, 259, 159
———. 1984, MNRAS, 209, 543
———. 1993, ApJ, 417, 351
———. 1995, ApJ, 439, 224
———. 1996, ApJ, 469, 906
———. 1997a, ApJ, 483, 309
———. 1997b, Science, 276, 1836
———. 1998a, Ann. Rev. Earth Planet. Sci., 26, 53
———. 1998b, Nature, 393, 141
Boss, A. P., & Myhill, E. A. 1992, ApJS, 83, 311
Butler, R. P., & Marcy, G. W. 1996, ApJ, 464, L153
Butler, R. P., Marcy, G. W., Williams, E., Hauser, H., & Shirts, P. 1997, ApJ, 474, L115
Cameron, A. G. W. 1978, Moon Planets, 18, 5
———. 1995, Meteoritics, 30, 133
Cameron, A. G. W., & Truran, J. W. 1977, Icarus, 30, 447
Cassen, P. M., Smith, B. F., Miller, R. H., & Reynolds, R. T. 1981, Icarus, 48, 377
Chabrier, G., Saumon, D., Hubbard, W. B., & Lunine, J. I. 1992, ApJ, 391, 817
Chandrasekhar, S. 1961, Hydrodynamic and Hydromagnetic Stability (New York: Dover)
Cochran, W. D., Hatzes, A. P., Butler, R. P., & Marcy, G. W. 1997, ApJ, 483, 457
DeCampli, W. M., & Cameron, A. G. W. 1979, Icarus, 38, 367
Drimmel, R. 1996, MNRAS, 282, 982
Gatewood, G. 1996, BAAS, 28, 885
Guillot, T., Chabrier, G., Morel, P., & Gautier, D. 1994, Icarus, 112, 354
Guillot, T., Gautier, D., & Hubbard, W. B. 1997, Icarus, 130, 534
Heemskerk, M. H. M., Papaloizou, J. C., & Savonije, G. J. 1992, A&A, 260, 161
Holman, M., Touma, J., & Tremaine, S. 1997, Nature, 386, 254
Kuiper, G. P. 1951, Proc. Natl. Acad. Sci., 37, 1
Laughlin, G., & Bodenheimer, P. 1994, ApJ, 436, 335
Laughlin, G., Korchagin, V., & Adams, F. C. 1997, ApJ, 477, 410
Laughlin, G., & Różyczka, M. 1996, ApJ, 456, 279
Lay, O. P., Carlstrom, J. E., & Hills, R. E. 1997, ApJ, 489, 917
Lin, D. N. C., & Papaloizou, J. 1986, ApJ, 309, 846
Lissauer, J. J. 1987, Icarus, 69, 249
Marcy, G. W., Butler, R. P., Williams, E., Bildsten, L., Graham, J. R., Ghez, A. M., & Jernigan, J. G. 1997, ApJ, 481, 926
Mayor, M., & Queloz, D. 1995, Nature, 378, 355
Mazeh, T., Krymolowski, Y., & Rosenfeld, G. 1997, ApJ, 477, L103
Mizuno, H. 1980, Prog. Theor. Phys., 64, 544
Mumma, M. J. 1996, Nature, 383, 581
Nelson, A. F., Benz, W., Adams, F. C., & Arnett, D. 1998, ApJ, 502, 342
Noyes, R. W., Jha, S., Korzennik, S. G., Krockenberger, M., Nisenson, P., Brown, T. M., Kennelly, E. J., & Horner, S. D. 1997, ApJ, 483, L111
Ohashi, N., Hayashi, M., Ho, P. T. P., Momose, M., & Hirano, N. 1996, ApJ, 466, 957
Ostriker, E. C., Shu, F. H., & Adams, F. C. 1992, ApJ, 399, 192
Pickett, B. K., Durisen, R. H., & Davis, G. A. 1996, ApJ, 458, 714
Pickett, B. K., Durisen, R. H., & Link, R. 1997, Icarus, 126, 243
Pollack, J. B. 1984, ARA&A, 22, 389
Pollack, J. B., Hubickyj, O., Bodenheimer, P., Lissauer, J. J., Podolak, M., & Greenzweig, Y. 1996, Icarus, 124, 62
Shu, F. H., Tremaine, S., Adams, F. C., & Ruden, S. P. 1990, ApJ, 358, 495
Slattery, W. L., DeCampli, W. M., & Cameron, A. G. W. 1980, Moon Planets, 23, 381
Stevenson, D. J. 1982, Planet. Space Sci., 30, 755
Stevenson, D. J., & Fishbein, E. 1981, in Lunar and Planetary Science XII (Houston: LPI), 1040
Strom, S. E., Edwards, S., & Skrutskie, M. F. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine (Tucson: Univ. Arizona Press), 837
Tomley, L., Cassen, P., & Steiman-Cameron, T. 1991, ApJ, 382, 530
Tomley, L., Steiman-Cameron, T. Y., & Cassen, P. 1994, ApJ, 422, 850
Toomre, A. 1964, ApJ, 139, 1217
Truelove, J. K., Klein, R. I., McKee, C. F., Holliman, J. H., Howell, L. H., & Greenough, J. A. 1997, ApJ, 489, L179
van Leer, B. 1977, J. Comput. Phys., 23, 263
Velusamy, T., Kuiper, T. B. H., & Langer, W. D. 1995, ApJ, 451, L75
Walker, G. A. H., Walker, A. R., Irwin, A. W., Larson, A. M., Yang, S. L. S., & Richardson, D. C. 1995, Icarus, 116, 359
Ward, W. R., & Hourigan, K. 1989, ApJ, 347, 490
Weidenschilling, S. J. 1977, ApSS, 51, 153
———. 1988, in Meteorites and the Early solar System, ed. J. F. Kerridge & M. S. Matthews (Tucson: Univ. Arizona Press), 348
———. 1997, in Lunar and Planetary Science XXVIII, ed. D. Blanchard & D. Black (Houston: LPI), 1513
Wetherill, G. W. 1990, Ann. Rev. Earth Planet. Sci., 18, 205
Yorke, H. W., Bodenheimer, P., & Laughlin, G. 1993, ApJ, 411, 274